ITEM 20 : INTERPRETATION D'UNE ENQUETE EPIDEMIOLOGIQUE

Epidémiologie = étude de la distribution des phénomènes de santé et de leurs déterminants
- Enquêtes épidémiologiques classées selon : la **procédure employée** (observation ou expérimentation), la **chronologie** (prospective, rétrospective ou transversale), le **choix des sujets** (en population, sur échantillon, exposé-non exposé, cas-témoin) et les **objectifs** (descriptif, analytique, évaluatif). Parfois hybride dans des enquêtes difficilement classables.
- **Etude expérimentale** = l'attribution de l'exposition est contrôlée par l'investigateur, de façon randomisée ou non → réalisable seulement si l'hypothèse testée est un effet bénéfique : efficacité d'un médicament (= essai thérapeutique), efficacité d'une stratégie médicale (diagnostique, thérapeutique ou préventive)
- **Etude quasi-expérimentale** = évaluation de programmes de santé publique (dépistage…) ou de pratiques médicales, dont l'allocation ne peut être randomisée : étude avant/après, étude ici/ailleurs, étude avant/après avec groupe témoin
- **Etude observationnelle** = observer la survenue d'un phénomène (maladie ou exposition) dans une population → seules études permettant de tester le rôle potentiellement délétère d'une exposition pour des raisons éthiques

Types d'enquête

Enquête observationnelle :
- **Enquête descriptive** : **transversale,** ou **longitudinale** (dont cohorte) → calcul de prévalence, incidence, mortalité…
- **Enquête analytique** : **transversale, cas/témoin, cohorte** (exhaustive, échantillon représentatif, exposé-non exposé) → analyse de facteurs de risques

Enquête expérimentale : descriptive (phase I), ou comparative (essai clinique, randomisé ou non)

Enquête descriptive

= Enquête visant à déterminer la fréquence, la variation et la répartition d'un phénomène de santé ou de leurs déterminants dans une population : habituellement simple et peu coûteuse
- Source de données : **information recueillie de manière systématique** (registre de maladie, réseaux de veille, enregistrement des causes de décès, maladies à déclaration obligatoire…) ou **enquêtes spécifiques**
- Permet de formuler des hypothèses sur des facteurs de risque d'une maladie selon sa répartition (temps/espace)

Enquête transversale *(cross-sectional)*

= Vision instantanée d'une situation épidémiologique, sans suivi des individus
- Sur l'ensemble d'une population (**enquête exhaustive**) ou **sur un échantillon**
- Peut être répétée dans le temps ou l'espace : variations géographiques/temporelles

→ Etude de **prévalence**

Enquête de cohorte descriptive

= Ensemble de sujets, ayant initialement une **caractéristique commune** (exposition…), indemne de l'événement étudié au cours de l'inclusion, **suivis au cours du temps**
- Sur l'ensemble d'une population (**enquête exhaustive**) ou **sur un échantillon**
- **Prospective** (information sur l'exposition recueillie avant que l'événement de santé soit survenu chez les sujets)**,** ou **rétrospective (historique** : information sur l'exposition recueillie après que l'événement de santé est survenu chez une partie au moins des sujets)

→ Etude de **l'incidence** (**enquête d'incidence**), de la **mortalité ou létalité** (**étude de survie**)

Enquête analytique

= Enquête visant à déterminer s'il existe un **lien statistique entre une exposition et une maladie**, toujours comparative (≥ 2 groupes) → ne détermine pas la causalité du lien. Permet aussi de chercher les facteurs associés aux issues défavorables (études pronostiques : mortalité, récidive, complications…) ou à la réponse au traitement.
- **Prospective** ou **rétrospective**, **fixe** (inclusion en début de période) ou **dynamique** (inclusion au fur et à mesure)

Enquête de cohorte *(cohort study)*

= Comparaison d'une incidence chez les sujets exposés et les sujets non exposés
- Les sujets doivent être indemnes de l'événement étudié à l'inclusion
- Le suivi des sujets non exposés doit être strictement similaire à celui des sujets exposés

Enquête de cohorte sur échantillon ou en population

= Mesure de l'exposition à l'inclusion (parfois répété lors du suivi) et enregistrement de la survenue éventuelle de la maladie au cours du suivi : estimation de la **prévalence de l'exposition à l'inclusion** et de **l'incidence de la maladie pendant le suivi**
- Echantillon : sujets extraits par sondage de la population source → résultats généralisables à la population cible si l'échantillon est représentatif
- Ensemble de la population cible = enquête exhaustive

→ Ne peut être réalisé que pour une exposition et une maladie fréquentes

Enquête exposé-non exposé

= Suivi de 2 cohortes, l'une exposée et l'autre non exposée, à la recherche de la survenue de l'événement étudié
- Le nombre de sujets exposés et non exposés est fixé a priori → l'incidence de la maladie ne peut pas être estimée directement
- Ne permet pas l'étude de plusieurs expositions simultanément

→ Privilégié pour une exposition rare et un événement relativement fréquent (milieu professionnel : amiante, solvant…)

Types d'enquête	*Enquête analytique*	**Enquête cas-témoin** *(case-control study)*	= Recherche dans le passé de 2 groupes de sujets, l'un atteint de la maladie (cas) et l'autre indemne (témoins), l'exposition antérieure à un ou plusieurs facteurs de risque potentiels - Sélection des cas et des témoins selon la maladie, en insu de l'exposition - Les cas doivent être représentatif de la population de malades et les témoins de la population d'où sont issus les cas, les 2 groupes doivent être comparables → **appariement** - Augmentation de la puissance si plusieurs témoins/cas, jusqu'à 4 témoins /cas - Recrutement de **cas prévalent** (présente la maladie au moment de l'enquête) ou **incident** (survenue de la maladie après la mise en place de l'enquête : recrutement prospectif) - Mais recueil de l'information **toujours rétrospectif** - Effectifs des cas et témoins fixés a priori → aucun calcul de prévalence ou d'incidence
		Transversale	- Parfois utilisée à visée analytique mais séquence chronologique impossible à déterminer
		Cas particuliers — **Enquête cas-témoins nichée**	= Cas et témoins recrutés parmi les **sujets d'une cohorte** suivie *(nested case-control study)* - Cas = tous les malades et témoins = tirés au sort parmi les non malades - Le **recueil prospectif des données** permet une meilleure qualité - **Respecte la séquence chronologique**
		Cas particuliers — **Enquête cas-croisé** *(case cross-over study)*	= Chaque cas est son propre témoin : comparaison de l'exposition lors de la période proche de l'événement étudiée (période cas) et de la période éloignée de l'événement (période témoin) - Analyse d'une **exposition ponctuelle intermittente** sur la **survenue d'un événement aigu** - Limite les biais et facteurs de confusion liés à la sélection des témoins - Utilisé pour étudier la relation entre la pollution atmosphérique et des maladies/mortalité
		Cas particuliers — **Enquête cas-cohorte** *(case cohort study)*	= Sélection aléatoire d'un échantillon de « **sous-cohorte** » lors de l'inclusion dans une cohorte, indépendamment de l'apparition de la maladie : avantages des cohortes et cas-témoins - Recueil des **cas incidents de la cohorte** et de la **sous-cohorte** - Analyse des **cas incidents** (sous-cohorte et non sous-cohorte) vs **non-cas de la sous-cohorte** - Avantage : - Etude de plusieurs expositions - Sous-cohorte représentative de la cohorte - Plus économique qu'une enquête de cohorte classique
	→ Guideline reports : . **STROBE** (**ST**engthening the **R**eporting of **Ob**servational studies in **E**pidemiology): études observationnelles . **CONSORT** (**CON**solidated **S**tandards **O**f **R**eporting **T**rials): essais contrôlés randomisés . **STARD** (**STA**ndards for the **R**eporting of **D**iagnostic accuracy studies) : études diagnostiques . **PRISMA** (**P**referred **R**eporting **I**tems for **S**ystematic reviews and **M**eta-**A**nalyses): revues systématiques, méta-analyses		

Type d'enquête	Sujets recrutés	Information recueillie	Mode de recrutement	Suivi des sujets	Recueil de l'information	Objectifs	Mesure d'association
Cohorte	Exposés Non-exposés	Maladie(s)	- Prospectif - Rétrospectif : cohorte historique - Transversal - Mixte	Oui	**Toujours prospectif**	- Incidence - Mortalité - Facteurs de risque ou pronostiques	- RR - OR
Cas-témoins	Malades Non-malades	Exposition(s)	- Incidents - Prévalents - Mixte	Non	**Toujours rétrospectif**	- Facteurs de risque ou pronostiques	- OR
Transversale	Tous	Maladie(s) Exposition(s)	Au début de l'enquête : T0	Non	À T0	- Prévalence - Facteurs de risque ou pronostiques	- RR - OR
Cas-témoins nichée	Malades Non-malades	Exposition(s)	Rétrospectif	Oui	Prospectif	- Facteurs de risque ou pronostiques	- OR
Cas-croisés	Malades		Prospectif				
Cas-cohorte	Cas Non-cas		Prospectif				- RR - OR

Etude	Cohorte prospective	Exposés-non exposés	Cas-témoins
Généralités	- **Longitudinale** (= sur une période), **prospective** - Niveau de preuve HAS : 2 → Incidence (pour la cohorte prospective), risque relatif (RR), nombre de personnes-années		- **Transversale, rétrospective** - Niveau de preuve HAS : 3 → Prévalence, odds ratio (OR)
Avantages	- Mesure de **l'incidence, du RR** (si échantillon représentatif) - **Recueil fiable des expositions** - Séquence chronologique entre exposition et maladie - Etude de plusieurs maladies et plusieurs facteurs de risque	- Mesure du RR - Recueil fiable des expositions - Séquence chronologique - Etude de plusieurs maladies - Adapté aux expositions rares	- Adapté aux maladies rares et de latence longue entre exposition et maladie - Etude de plusieurs facteurs de risque - Moindre coût, résultats plus rapides
Inconvénients	- Risque de perdus de vue - Peu adapté si exposition rare, maladies rares, temps de latence long entre exposition et maladie - Ethique - Long, coût élevé	- Pas de calcul d'incidence - Peu adapté si temps de latence long entre exposition et maladie - Ethique - Long, coût élevé	- Séquence chronologique dure à établir - Pas de calcul d'incidence ou RR (Estimation indirecte par l'OR) - Non adapté à l'étude de plusieurs maladies, si exposition rare - Risque de biais de classement (recueil rétrospectif)

Jugement de signification statistique

- Erreur aléatoire : résulte des **fluctuations d'échantillonnage** (hasard)
- Enquête analytique pour tester l'existence d'associations entre une (des) exposition(s) et un (des) événement(s) :
. **Univariée** : test du Chi2 ou probabilité exacte de Fisher (%), test de Student (moyennes), tests non paramétriques
. **Multivariée** (ajuste sur les facteurs de confusion potentiels) : régression logistique, linéaire, modèle de Cox. Des tests d'interaction déterminent si le risque d'événement lié à un facteur diffère selon l'exposition à un autre
→ Le degré de signification statistique reflète la confiance que l'on peut accorder aux résultats : interprétable ou non

Risque α	= **Risque de 1ère espèce** : risque de conclure à une différence alors qu'il n'y en a pas (faux positif) - Arbitrairement fixé à **< 5%** (seuil de significativité)
Risque β	= **Risque de 2nd espèce** : risque de ne pas conclure à une différence alors qu'elle existe (faux négatif)
Puissance Δ	= Probabilité de mettre en évidence une différence qui existe réellement : **Δ = 1 – β**
p value	= Degré de significativité d'une différence observée : probabilité que le résultat soit lié au hasard - Si **p < 0,05** : association significative - Si **p > 0,05** : manque de puissance (la différence peut être due au hasard) → on ne peut pas conclure → La p-value est le calcul *a posteriori* de la significativité des résultats obtenus, alors que le risque α est une estimation *a priori* de la puissance voulue. Surtout important dans les études expérimentales car les études observationnelles cherchent plutôt à estimer l'ampleur des risques : intervalle de confiance
Intervalle de confiance	= Fluctuation des résultats pour la même étude dans les mêmes conditions : si l'on répète l'étude, le résultat observé se retrouvera dans 95% des cas dans l'intervalle de confiance

Calcul des risques

- Risque = probabilité qu'un événement survienne dans une population pendant une période donnée. Estimé par la fréquence de l'événement dans l'échantillon

Tableau 2x2

	Exposé	Non exposé
Malade	**a**	**b**
Non malade	**c**	**d**

Risque en excès (excès de risque)

= Différence De Risque : modèle additif, différence entre le risque absolu et le risque de base

$$\textbf{DDR} = \textbf{incidence}_{\text{exposé}} - \textbf{incidence}_{\text{non exposé}} = \frac{A}{A+C} - \frac{B}{B+D}$$

- Correspond à la Réduction Absolue du Risque (RAR) utilisée dans les essais
- L'absence de relation entre la maladie et le facteur correspond à un excès de risque de 0

Risque relatif

= Modèle multiplicatif, rapport entre les incidences de maladies chez les exposés et les non exposés

$$\textbf{RR} = \frac{A\ /\ (A+C)}{B\ /\ (B+D)}$$

- Calculé seulement dans des **enquêtes prospectives**
- Interprétation : - RR > 1 : excès de risque dans la population exposée = **facteur de risque**
- RR < 1 : diminution du risque dans la population exposée = **facteur protecteur**
- RR = 1 : l'exposition n'a pas de conséquence sur l'apparition de la maladie

→ **Réduction relative du risque** = **1 – RR**, augmentation relative du risque = RR – 1

Calcul des risques

Odds ratio

= Rapport de cotes : dans les **enquêtes rétrospectives** (Ø incidence : risque relatif non calculable)

$$OR = \frac{\text{Odds exposés}}{\text{Odds non exposés}} = \frac{A \times D}{B \times C}$$

- **Approximation du risque relatif qu'il surestime, d'autant moins que la maladie est rare**

Mesures d'impact

Reflètent l'importance du risque dans la population :

- **Risque attribuable = fraction de risque attribuable** : proportion de cas imputable au facteur

$$RA = \frac{\text{fréquence de l'exposition} \times (RR-1)}{\text{Fréquence de l'exposition} \times (RR-1) + 1}$$

Dans une enquête cas-témoins estimation de RR par l'OR et de la fréquence d'exposition par la prévalence d'exposition chez les témoins.

➝ N'a de sens que si la relation causale est probable. Permet d'estimer le nombre de cas évitables.

- **Fraction étiologique** : proportion de cas attribuables au facteur parmi les sujets exposés :

$$FE = \frac{RR - 1}{RR}$$

Biais

= **Erreur systématique** dans l'estimation d'un paramètre

- Susceptible de masquer, renforcer ou créer une liaison entre facteur et maladie ➝ sur- ou sous-estimation du risque
- Doivent être limités le plus possible dès la rédaction du protocole
- Ne peuvent pas être diminués en augmentant la taille de l'échantillon

Biais de sélection

= Erreur systématique dans l'estimation d'un paramètre due au mode de constitution de l'échantillon ou au suivi des sujets ➝ absence de représentativité de la population cible

- De recrutement, d'échantillonnage et d'inclusion :

. Enquêtes descriptives: recrutement hospitalier, critères d'inclusion trop sélectifs, défaut de participation

. Enquêtes transversales: absence de sujets liée à la maladie ou l'exposition analysée

. Enquêtes comparatives: biais différentiel si sélection non indépendante de l'exposition (cas-témoins) ou du risque d'événement (cohorte)

- D'attrition = perdus de vue

➝ Doit être prévenu lors de la conception du protocole de l'étude (ne peut être contrôlé *a posteriori*) : sélection des cas et témoins sans connaître leurs expositions, minimiser les perdus de vue et les analyser…

Biais de mesure = classement = information

= Erreur systématique dans la mesure de l'exposition ou la détermination de la maladie : mauvais classement des sujets en exposés/non exposés ou malade/non malade

- **Non différentiel** : l'erreur de mesure ne diffère pas entre les groupes (instrument mal étalonné…)

➝ Sous-estime la force de l'association maladie-exposition

- **Différentiel** = l'erreur de mesure diffère entre les groupes : **biais d'investigation, d'analyse** ➝ évité par une procédure en aveugle, la formation des enquêteurs, un interrogatoire standardisé…

➝ Sous-estime ou plus souvent surestime la force de l'association maladie-exposition

Facteur de confusion

= Interférence d'un facteur dans la relation de causalité, lié à la fois à la maladie et au facteur de risque

Méthode de réduction	
	- Appariement : formation de paires en associant à chaque cas à un ou plusieurs témoins de mêmes caractéristiques (âge, sexe…) ➝ améliore la comparabilité **- Analyse stratifiée** et/ou **multivariée** : analyse de l'effet d'une exposition en prenant en compte les effets des autres facteurs . **Stratification** : répartition des sujets en strates selon le facteur ➝ si le RR ou l'OR n'est pas comparable entre les strates, il s'agit d'un facteur d'interaction . **Test multivarié** : prise en compte d'un facteur de confusion a posteriori

Causalité

Critères de causalité de Bradford-Hill

- Relation causale n'est discutée qu'en cas d'augmentation significative du risque sans biais majeurs
- Critères de causalité des enquêtes observationnelles de Bradford-Hill (1965), faisceau d'arguments :

Critères internes à l'étude	Critères externes à l'étude
- **Séquence dans le temps** : exposition précède la maladie - **Force de l'association** - **Spécificité de l'association** : exposition présente chez tous les malades ou presque et seulement chez eux, et pour cette seule maladie - **Relation dose-effet** - **Cohérence interne** : méthodologie, prise en compte des biais	- **Constance de l'association** : association constante et reproductible dans diverses situations - **Plausibilité biologique** : existence d'un mécanisme d'action plausible - **Cohérence externe** : reproductibilité des résultats - **Preuve expérimentale** : expérimentation animale, essai expérimental (une intervention visant à diminuer le FdR diminue la maladie) - **Analogie** : avec une exposition et maladie analogues